DOCUMENT RESUME

ED 365 711 TM 020 902

AUTHOR	Inger, Morton
TITLE	Authentic Assessment in Secondary Education.
INSTITUTION	Columbia Univ., New York, NY. Inst. on Education and the Economy.
SPONS AGENCY	William T. Grant Foundation, Washington, DC. Commission on Work, Family, and Citizenship.
REPORT NO	ISSN-1059-2776
PUB DATE	Mar 93
NOTE	5p.; This Brief is drawn from a report by Clifford Hill and Eric Larsen, entitled "Testing and Assessment in Secondary Education: A Critical Review of Emerging Practices." The research was conducted with support from the National Center for Research in Vocational Education (NCRVE), funded by the Office of Vocational Education, U.S. Department of Education.
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Collected Works - Serials (022)
JOURNAL CIT	IEE Brief; n6 Mar 1993
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Constructed Response; Curriculum Development; *Educational Assessment; Educational Change; Educational Objectives; Educational Practices; Equal Education; Models; *Portfolios (Background Materials); School Administration; Secondary Education; *Student Evaluation; Test Construction; *Test Use
IDENTIFIERS	Alternative Assessment; *Authentic Assessment; Performance Based Evaluation

ABSTRACT

One educational change that is beginning to take hold is the use of alternative ways of testing and assessment, often referred to as authentic assessment. Advocates of authentic assessment believe that conventional testing is distorting educational goals, while authentic assessment can foster good educational practices. In fact, many educators view authentic assessment as an enterprise altogether different from testing. An extended time frame for authentic assessment allows tracking student involvement with multiple tasks in documentation practices such as portfolios and exhibitions. In support of authentic assessment, advocates argue for its excellence and support the use of constructed responses and direct assessment in pursuit of greater equity and efficiency. The goals of authentic assessment are: (1) reforming curriculum and instruction; (2) improving teacher morale and performance; and (3) strengthening student commitment and capacity for self-monitoring. It must be recognized that authentic assessment makes far greater demands on both teacher and student than conventional assessment, and that these demands require massive administrative support. (Contains 10 references.) (SLD)

ED 365 711

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

INSTITUTE ON EDUCATION AND THE ECONOMY

IEE BRIEF

ISSN 1059 2776

Number 6/March 1993

AUTHENTIC ASSESSMENT IN SECONDARY EDUCATION

The remarkable consensus that our educational system is not working as well as it should has led to many calls for the restructuring and reforming of American education. One reform that is beginning to take hold is the adoption of alternative ways of testing and assessment, often referred to as authentic assessment. Conventional testing was designed primarily to provide feedback on how well specific knowledge and skills have been learned. Advocates of *authentic assessment* see this as an important function, but they believe that how and what we test has a powerful influence on how and what is taught. They believe that conventional testing is distorting educational goals, whereas authentic assessment can foster good educational practices.

Models of Education

Conventional testing operates with standards arising almost exclusively from the education system itself. Life inside schools is divided into self-contained compartments such as math and history, and what counts as desirable within each of these compartments is largely generated from within. Outside of school, of course, life is not divided into separate subject-matter categories. By staying almost exclusively within the boundaries of this school model, conventional testing reinforces the school's separation from the world outside of school.

Berryman and Bailey (1992) have described a fundamental dualism that permeates education, a dualism

> between culture and vocation, head and hand, abstract and concrete, theoretical and applied (p. 106).

To dissolve this dualism, advocates of authentic assessment want assessments, and, as a consequence, schools, to focus on the broad knowledge and skills that individuals need to solve real-world problems. For example, students would not be assessed separately for mathematical problem-solving and for writing skills. They would deal with problems that call for them to use mathematical reasoning within a discursive structure. Moreover, they would be motivated to go beyond the purely academic *knowing* to the technical and social ways of *doing*. To expand the metaphor introduced by Berryman and Bailey, head would be brought together with both hand and heart.

Further, advocates of authentic assessment believe that schooling — and testing in particular — has been excessively concerned with language as a means of exhibiting what has been learned. Hence, authentic assessment requires students to actually produce things (e.g., an architectural model or a computer data base) or carry out activities (e.g., a science experiment or a survey in their community), not just talk or write about them.

Documenting Student Work

Many educators have come to view authentic assessment as an enterprise altogether different from testing. Testing, no matter how it is reformed, still focuses on how students handle tasks on a single occasion under severe time constraints. Such constraints lead to limited tasks that cannot get at the complexity of what people do when they engage in purposeful activity over a period of time. An extended time frame offers the opportunity for students to work on a greater range of tasks. Assessments that track student involvement with multiple tasks over time are called *documentation practices*.

Portfolios and exhibitions. Portfolios are the most widely used form of documentation. Teachers have long used portfolios to help students keep track of their work and present it in an organized way. What is new is the notion that portfolios can replace, or at least supplement, conventional testing.

Portfolios used in this way differ from traditional classroom-oriented portfolios in at least three ways: (1) the items included tend to be prescribed and can even include timed tasks; (2) there is more insistence on individual as opposed to collaborative work; (3) and evaluation schemes tend to be analytic — breaking the work down into component parts and features and analyzing each separately — rather than holistic — giving an assessment of the work as a whole.

Exhibitions are displays of personal creations, artifacts, or performances. They are especially used to assess student work in the sciences, the arts, and vocational education.

Records. Teacher-maintained records have a long tradition, and proponents of authentic assessment are encouraging teachers to be more observant and more responsive to individual student characteristics. Some assessment models emphasize student-maintained records as an ideal way of encouraging students to assess their own strengths and weaknesses. As students engage in keeping records of their own work, they develop skills that are crucial to success in school and the workplace: the capacity to organize information and store it in such a way that they can easily retrieve it. In effect, documentation practices teach the very skills they are designed to assess.

Assessing Authentic Assessment

In the early part of this century, educators such as Thorndike (1913) invoked the principles of excellence, equity, and efficiency to support the development of multiple-choice testing. These same principles are now invoked to make the case for authentic assessment.

Excellence

For those who advocate authentic assessment, the pursuit of excellence is at the heart of their policies. They argue that such assessment, in contrast to conventional testing,

- requires students to construct responses rather than select among a menu of options

- elicits from students higher-order thinking in addition to basic low-level skills
- is designed to assess performance directly rather than rely on indirect measures
- is integrated with classroom instruction rather than separated from it.

Constructed responses. Conventional testing, as exemplified by multiple-choice tests, forces students to select from a given set of options. In addition to placing students into a passive posture, such forced selection can encourage the use of mechanical procedures in which students attend primarily to surface format; they frequently, for example, select an option simply because it differs from the others in length or structure.

When students are not given a menu of answers and must construct their own responses, they are forced to take the more active stance that characterizes everyday problem-solving. In constructing a response, they must decide on what is relevant, organize it in some way, and then work out its presentation.

From the standpoint of motivation, constructed-response tasks offer a clear advantage over conventional testing. Although students prefer multiple-choice tests (because they are "easier"), they typically spend more time preparing for a test with constructed-response tasks than they do for multiple-choice tests (Warren, 1979; D'Ydewalle, Swerts, & De Corte, 1983; Traub & MacRury, 1990). Documentation practices motivate students to work at the highest level: they invest greater energy and achieve higher standards.

Higher-order thinking. Since multiple-choice tests require tasks with a clear right answer, it is difficult, if not impossible, to construct tasks that require the complex thinking needed in real problem-solving.

Advocates of authentic assessment place major emphasis on developing tasks that elicit higher-order thinking (Resnick, 1987; Haney & Madaus, 1989; Newmann, 1991). They believe such tasks should be a central focus in alternative tests as well as in documentation practices. Thus they construct essay tests around real-world problems that do not have a simple solution, for example, how to protect the environment in a modern economy.

The concern with higher-order thinking, however, can introduce problems of its own. It can lead to an unwarranted opposition between basic skills and higher-order thinking. Basic skills are embedded in, not opposed to, higher-order thinking skills. For example, students are often encouraged to compose directly at the computer, but such composition is impossible without keyboarding skills. Authentic assessment must not bypass basic skills but rather find ways of evaluating such skills in relation to the higher-order thinking that they support.

Direct assessment. Conventional testing proceeds by indirection, presenting a range of tasks that sample discrete bits of knowledge or skill. In contrast, authentic assessment is committed to assessing student knowledge and skill more directly. But we must raise a caution flag here. Direct assessment of student knowledge and skill is difficult to accomplish within a test, no matter how much it is reformed. The time limits of a test constrain what students can do, and a test radically alters the context. Even if it is an authentic task, students carry it out — *in a test* — simply to display certain capacities that need to be assessed.

That problem is avoided when students carry out projects apart from a testing situation, for example, documentation projects requiring extended discourse, the creation of complex artifacts, and multifaceted performances. For example, when material from the workplace is introduced into schools, it is more likely to be effective — and provide opportunities for direct assessment of student skills — when it is used in an extended project rather than in a limited task on a test.

Integration with classroom instruction. A major criticism of conventional testing is that it undermines what goes on in the classroom. This is particularly true of a high-stakes test like the SAT. Because teachers feel obligated to prepare students for a test with such important consequences, they often spend an inordinate amount of class time on test-taking techniques. Even teachers committed to the development of higher-order thinking skills often spend too much time on low-level skills that are not connected to their larger goals.

Advocates of authentic assessment believe that assessment practices, if sufficiently aligned with curriculum and instruction, can become a powerful means of achieving excellence in the classroom. If these practices were adopted, "teaching to the test" would no longer have negative connotations; testing would become a further resource for developing classroom instruction.

Equity

Advocates claim that authentic assessment will lead to greater equity because, in contrast to conventional testing, it

- uses multiple samples of student work collected over an extended period of time
- is based on clear criteria that students are made aware of
- allows for the possibility of multiple human judgments.

Multiple samples of extended work. Conventional testing is designed to be administered during a normal school period, and the strict time limit creates anxiety that prevents students from concentrating on what they must do. It also presents a series of discrete tasks that force students to move rapidly from one unconnected item to the next.

In contrast, authentic assessment based on extended projects rewards sustained attention to an extended task. As students engage in a task over an extended time, they are able to produce more substantial work. An extended project gives them a chance to demonstrate knowledge and skills that are bypassed in conventional testing. The result is a more equitable evaluation.

But authentic assessment introduces its own equity issue: how to determine how much help a student has received

in his or her project. To deal with this issue, students can be asked to document the development of their work — whether it be essays, tooled machine parts, or videotaped diagnostic procedures — with notes, prototypes, or early drafts or videotapes. In other instances, there is simply less concern that work samples be the product of a single individual. Indeed, a fundamental concern for many who advocate authentic assessment is to determine to what degree students can work effectively with others.

Public criteria. In the United States, conventional testing has long been shrouded in secrecy (Schwartz & Viator, 1990). The rationale for this policy is that the integrity of individual tests must be insured.

Advocates of authentic assessment are opposed to this secrecy. They claim that students should know how they are evaluated so they can adequately prepare for what is expected of them.

As admirable as a policy of public disclosure may be, it may not be as effective as anticipated. It is no easy matter to present a complex set of criteria so that a broad range of students can understand them and apply them to their work. There is also the danger that any explicit statement of criteria will not reflect how judgments are actually made.

Multiple human judgments. To insure strict impartiality, conventional tests are designed so that human judgment is removed from the evaluation of test responses. But human judgment is not removed; it is simply displaced to an earlier stage. Someone still has to decide which tasks are to be included on the test, how they will be weighted, which answers are correct, and so on. Because these decisions are buried in the process, no individuals can be held accountable (Hill & Parry, in press).

Authentic assessment, by contrast, is often designed so that students can appeal the evaluation they receive. Generally, two or more individuals are required to evaluate student work, and there is a moderating system to review the initial evaluation.

But problems are inevitable when assessment is carried out by teachers who work directly with the students. It is clear that teachers' personal relations with students affect the evaluations they make. Personal feelings are difficult to monitor since they often remain unconscious, particularly where they operate across gender and ethnic lines.

Efficiency

Proponents of conventional testing claim that the multiple-choice format facilitates both administration and scoring, making it possible to administer a single test in less than two hours to thousands of high school students nationwide and then score it and send it to hundreds of college and university admission offices in a matter of weeks—all at relatively low cost.

But those who support authentic assessment believe that the claims of efficiency for conventional testing are superficial and misleading. They argue that these claims mask the massive inefficiency that arises in the educational system when fundamental goals are distorted by inappropriate methods of assessment, for example, when teachers must spend inordinate amounts of time teaching students test-taking techniques. They thus argue for a deeper notion of efficiency, one that has to do with the degree to which assessment fosters good educational practices.

The Goals of Authentic Assessment

Authentic assessment has three fundamental goals:

1. reforming curriculum and instruction
2. improving teacher morale and performance
3. strengthening student commitment and capacity for self-monitoring.

1. Reforming curriculum and instruction. Since technology is continuously restructuring what people must do in the workplace, it is not useful for students to concentrate on a highly specific set of skills. They must know how to work with the information that technology produces and how to use that information to communicate with others.

In focusing on actual needs of the workplace, authentic assessment acclimates students to confronting basic questions they will face when they enter the workforce. Once authentic practices are in place, they can provide an effective way to insure that school curriculum and instruction reflect the changing needs of the larger society.

2. Improving teacher morale and performance. One of the major benefits of authentic assessment is that classroom teachers assume a central role in evaluating students. Traditionally, teachers performed this role, but with the advent of conventional testing, their responsibility and authority were greatly reduced. One consequence of this policy was that teacher morale was damaged. They came to feel that their own judgment about students was not trusted and that conventional testing was developed as a kind of surrogate. Authentic assessment has restored to teachers their right and responsibility to evaluate students, for most of these practices depend upon teacher judgment. Moreover, teachers are actively involved in the development of authentic practices. In many school districts throughout the country, teachers meet regularly to develop new assessment methods.

Authentic assessment calls for a clinically oriented style of observing student work. This method of assessment powerfully transforms what classroom teachers do: for example, they can help students plan their work; then, throughout the semester, they read drafts or view prototypes and make suggestions. As the semester nears an end, teachers help students select what to include in their portfolios; and once the portfolio is submitted, teachers review the work in preparation for a final conference.

3. Strengthening student commitment and capacity for self-monitoring. Advocates of authentic assessment question the

Institute on Education and the Economy
Box 174, Teachers College, Columbia University
New York, New York 10027
212-678-3091
Director: Thomas R. Bailey

widely accepted notion that external evaluation motivates students; external evaluation does not take sufficient account of the strong commitment that is engendered when students accept responsibility for their own work. Within authentic assessment, students are increasingly expected to select the projects on which they will be evaluated. Once they have made this choice, they are responsible for getting the necessary work done. Although teachers are available for consultation and coaching, students are expected to demonstrate self-reliance, which is, in fact, a crucial quality to be evaluated.

Moreover, students are increasingly expected to evaluate their own work. In one model, the individual student and teacher — or supervisor in a workplace apprenticeship — make independent evaluations of the student's work. Once these evaluations are in place, the student and teacher work together to negotiate a mutually acceptable assessment. This process helps students come to accept continuous assessment as a natural and valuable part in any work that they do. Such acceptance enables them to acknowledge their own limitations and yet strive to achieve work that makes full use of their knowledge and skills.

In the final analysis, the major goal of any assessment should be to develop in students the capacity and the commitment to monitor their own work. Students will not learn to produce work of excellent quality as long as standards remain external to them. It is only as they internalize standards that they are able to engage in the rigorous monitoring that insures excellent work.

Authentic assessment, when compared to conventional testing, makes far greater demands on both students and teachers. Some critics believe that it takes so much time that instruction is shortchanged. But this point of view misses the symbiotic relation between instruction and assessment. Within the best models of authentic assessment, teaching and evaluation become virtually indistinguishable: an assessment that teaches students how to monitor their work is a vital form of instruction.

It is important to recognize, however, that the intensive labor demands of authentic assessment require massive administrative support. The potential of authentic assessment as a force for reform in education is crucially dependent upon firm and continuous administrative support.

— Morton Inger

References

Berryman, S., & Bailey, T. (1992). *The double helix of education and the economy.* New York: Institute on Education and the Economy, Teachers College, Columbia University.

D'Ydewalle, G., Swerts, A., & De Corte, E. (1983). Study time and test performance as a function of test expectations. *Contemporary Educational Psychology, 8,* 52-59.

Haney, W., & Madaus, G. (1989). Searching for alternatives to standardized tests: Whys, whats, and whithers. *Phi Delta Kappan, 70*(9), 683-687.

Hill, C., & Parry, K. (In press). *Testing and assessment: International perspectives on English literacy.* Harlow, UK: Longman.

Newmann, F. (1991). Linking restructuring to authentic student achievement. *Phi Delta Kappan, 72*(2), 458-463.

Resnick, L. (1987). *Education and learning to think.* Washington, DC: National Academy Press.

Schwartz, J., & Viator, K. (1990). *The prices of secrecy: The social, intellectual, and psychological costs of assessment practice.* A Report to the Ford Foundation. Cambridge, MA: Educational Technology Center, Harvard Graduate School of Education.

Thorndike, E. (1913). *Educational psychology, Volume 1.* New York: Teachers College Press.

Traub, R., & MacRury, K. (1990). Multiple-choice vs. free-response in the testing of scholastic achievement. In K. Ingenkamp & R. Jager (Eds.), *Tests and trends 8: Jahrbuch der padagogischen diagnostik.* Weinheim and Basel, Germany: Beltz Verlag.

Warren, G. (1979). Essay versus multiple-choice tests. *Journal of Research in Science Teaching, 16*(6), 563-567.

This Brief is drawn from a report by Clifford Hill and Eric Larsen, entitled *Testing and Assessment in Secondary Education: A Critical Review of Emerging Practices.* The research was conducted with support from the National Center for Research in Vocational Education (NCRVE), funded by the Office of Vocational Education, U.S. Department of Education. To order the report, call the NCRVE publication department at 800-637-7652. Publication and dissemination of this Brief is supported, in part, by Teachers College, Columbia University, and the William T. Grant Foundation Commission on Work, Family, and Citizenship.